Calibrating uncertainty

80% confidence interval

A

B

10% chance the value is below the lower limit

80% chance the true value falls within the limits

10% chance the true value is above the upper limit

P10
Surprisingly low value

P90
Surprisingly high value
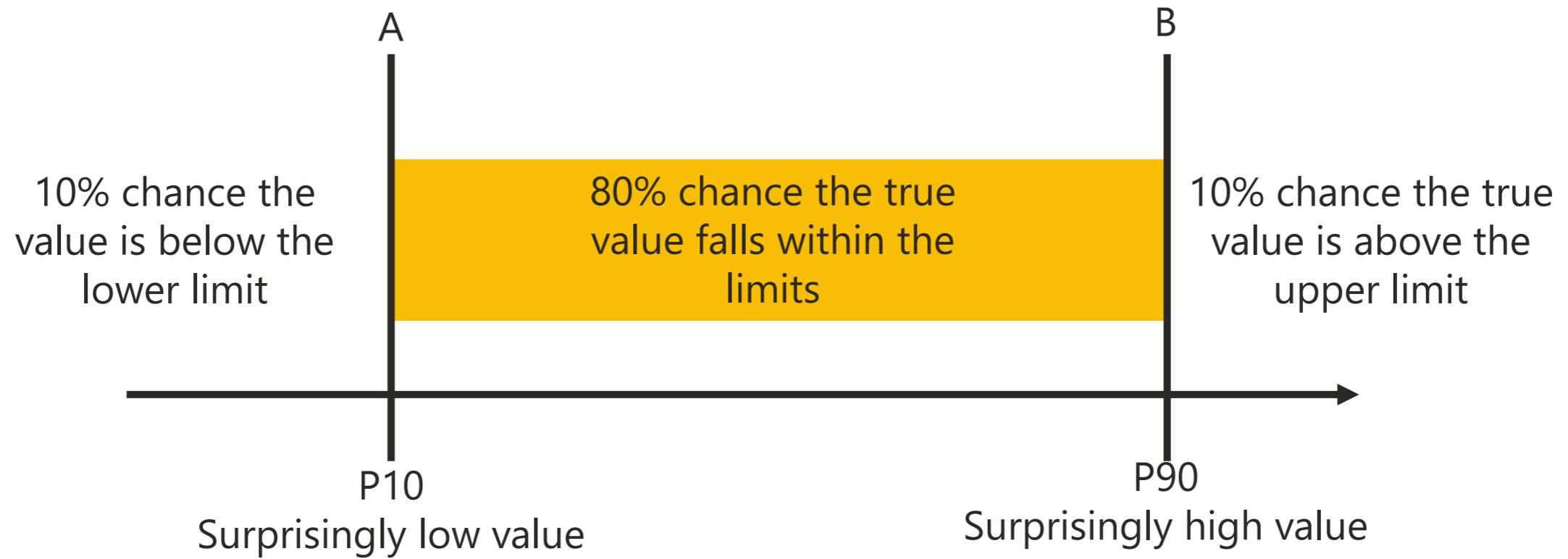
Exercise

Ten general knowledge questions.

For each question, provide a 80% confidence interval.

In your judgement, there should be

- a 10% chance that the true value lies below the lower limit

- a 10% chance that the true value lies above the upper limit

Don't agonise over it.

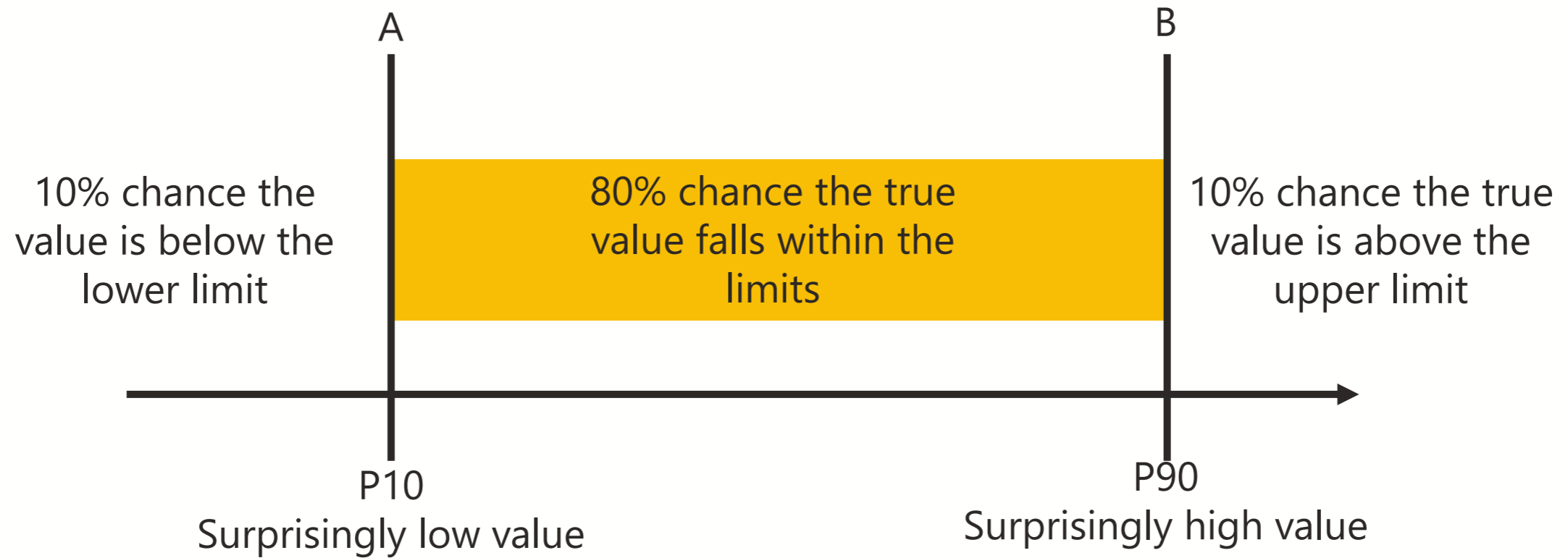**This is not a test of your general knowledge.**

You must give answers to each question.

		Lower bound 10% chance value is lower	Upper bound 10% chance value is higher	
1	When was Elvis Presley born?			
2	How many miles tall is Mount Everest?			
3	What percentage of the atmosphere is oxygen by weight?			
4	The first European printing press was invented in what year?			
5	What percentage of all electricity consumed by U.S. households in 2001 was used by kitchen appliances?			
6	What percentage of a square could be covered by a circle of the same width?			
7	In what year did Isaac Newton publish the universal laws of gravitation?			
8	At how many miles per hour can a peregrine falcon fly in a dive?			
9	How many calories are in 8 ounces of orange juice?			
10	How many people were permanently evacuated after the Chernobyl nuclear power plant accident?			

Exercise from 'How to measure anything', D Hubbard

Count the number of instances where the answer fell between your upper and lower bound	Lower bound 10% chance value is lower	Upper bound 10% chance value is higher	Answer
1	When was Elvis Presley born?			**1935**
2	How many miles tall is Mount Everest?			**5.5**
3	What percentage of the atmosphere is oxygen by weight?			**21%**
4	The first European printing press was invented in what year?			**1450**
5	What percentage of all electricity consumed by U.S. households in 2001 was used by kitchen appliances?			**26.7%**
6	What percentage of a square could be covered by a circle of the same width?			**78.5%**
7	In what year did Isaac Newton publish the universal laws of gravitation?			**1685**
8	At how many miles per hour can a peregrine falcon fly in a dive?			**200**
9	How many calories are in 8 ounces of orange juice?			**120**
10	How many people were permanently evacuated after the Chernobyl nuclear power plant accident?			**135,000**

80% confidence intervals

A

B

10% chance the value is below the lower limit

80% chance the true value falls within the limits

10% chance the true value is above the upper limit

P10
Surprisingly low value

P90
Surprisingly high value

A systematic approach to probabilistic forecasting p10-p50-p90

We are after realistic, plausible surprisingly low and surprisingly high forecast values not unrealistic, implausible or extremely surprising values.

We ask our experts to provide us with an interval that ranges from low to high (reflecting variability) with 80% degree of belief. This means that the expert is 80% confident the true value will lie within their interval 8 out of 10 times.

P10- the 10th percentile can be thought of as a "surprisingly low" value because there is a less than 1 in 10 chance of values below it.

P90- the 90th percentile can be thought of as a "surprisingly high" value because there is a less than 1 in 10 chance of values above it.

Extreme values stem from very rare events involving extreme conditions which are notoriously difficult to predict and are better handled by extreme event scenario planning (eg disaster planning).

The expert will first be asked to think carefully about the key factors which would lead to a surprisingly high value and list them out. These insights are an integral part of the process and are what leads to cogent, realistic and plausible forecasts

Repeat for surprisingly low values

(Only then, the most likely value - p50 (median))

Figure 1: The Combined Results of 11 Studies in Probability "Calibration" Training

Most experts significantly overstate their confidence in forecasts. Calibration training corrects this.

CALIBRATED PROBABILITY ASSESSMENTS
An Introduction

8
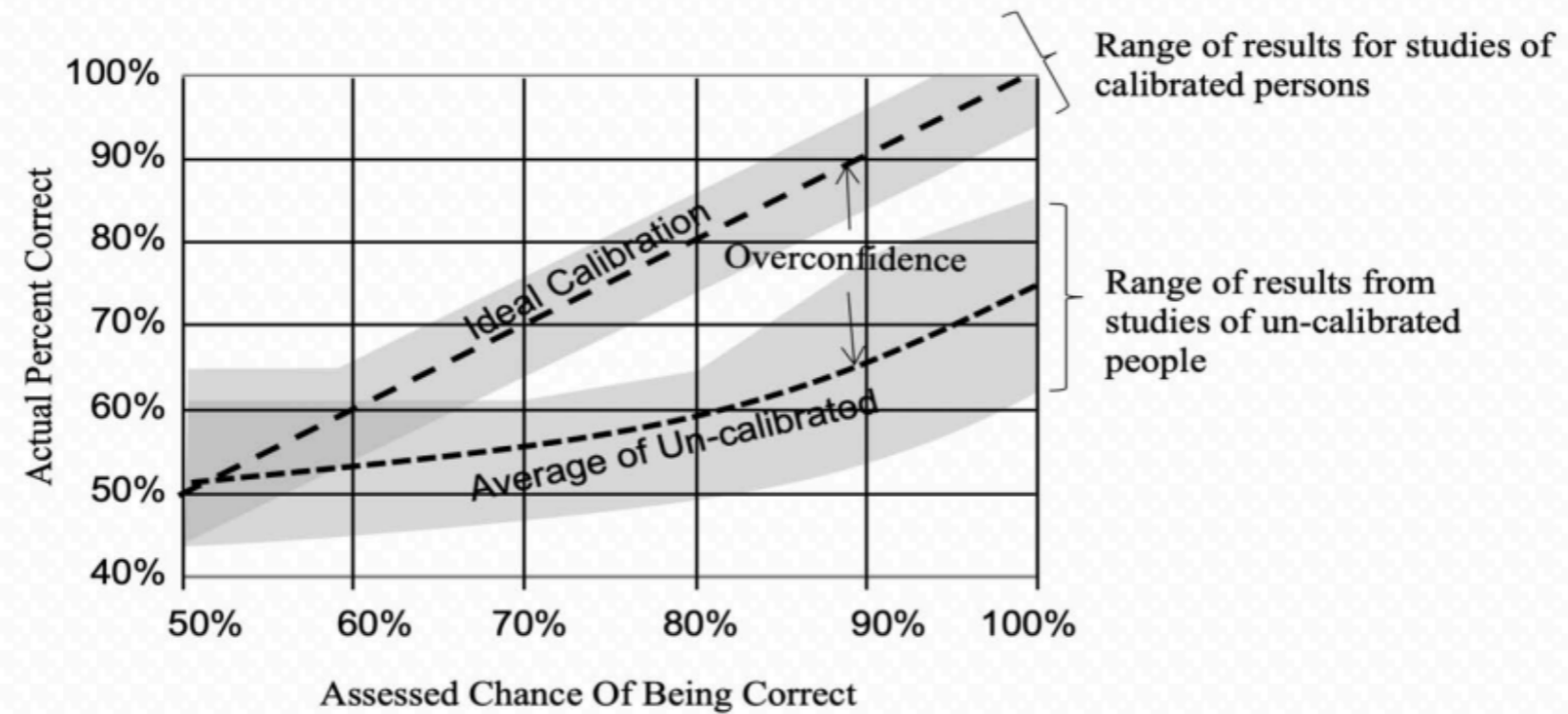
Reflections

If the intervals you gave **were** 80% confidence intervals, for how many questions would you expect (on average) the answer to fall within the intervals?

Why do you think that most people scored less than 8?

What can we take from this?

If you repeated the exercise, on a different set of questions, what might you do differently?